Atten Percept Psychophys (2011) 73:2528–2541
DOI 10.3758/s13414-011-0196-5

The effect of spatial competition between object-level representations of target and mask on object substitution masking

Duncan Guest · Angus Gellatly · Michael Pilling

Published online: 18 August 2011

Abstract One of the processes determining object substitution masking (OSM) is thought to be the spatial competition between independent object file representations of the target and mask (e.g., Kahan & Lichtman, 2006). In a series of experiments, we further examined how OSM is influenced by this spatial competition by manipulating the overlap between the surfaces created by the modal completion of the target (an outline square with a gap in one of its sides) and the mask (a four-dot mask). The results of these experiments demonstrate that increasing the spatial overlap between the surfaces of the target and mask increases OSM. Importantly, this effect is not caused by the mask interfering with the processing of the target features it overlaps. Overall, the data indicate, consistent with Kahan and Lichtman, that OSM can arise through competition between independent target and mask representations.

Keywords Visual perception · Attention

Introduction

Object substitution masking (OSM) refers to a phenomenon observed under a range of closely related but subtly varying experimental procedures in which a mask that trails the offset of a target impairs reporting of that target even though the two neither are physically superimposed nor contain closely abutting contours (Di Lollo, Enns, & Rensink, 2000; Enns & Di Lollo, 1997, 2000). In a typical example of substitution masking, known also as common onset masking, an array of stimuli is briefly presented, and one of the stimuli is simultaneously surrounded by a mask (usually consisting of four dots in a square formation [hereafter, 4DM]), which also serves as the cue to the target. In the standard paradigm, either the target and mask offset together (common-offset control condition) or the mask offsets sometime after the offset of the target (trailing mask condition). When the mask trails the target in this way, it interferes with processing of the target object, resulting in a substantial reduction in the ability to discriminate or identify the target (Di Lollo et al., 2000; Gellatly, Pilling, Cole, & Skarratt, 2006).

OSM has received considerable interest over the past 10 years because it is quite distinct from the types of masking that arise from low-level contour and feature interactions. Because OSM is a difference in performance between two conditions, in both of which the mask and target onset simultaneously, it cannot be attributed to the mask's interrupting early processing of the target (Turvey, 1973), since this must occur equally for the two conditions. Neither can OSM result from lateral inhibition processes in early vision, such as in metacontrast masking; unlike metacontrast masking, OSM is not influenced by distance between the adjacent contours of target and mask (Enns & Di Lollo, 1997), is completely eliminated when attention is focused on the target (Enns, 2004) as long as executive attention is not limited (Dux, Visser, Goodhew & Lipp, 2010), and is present when target and mask onset simultaneously (Bernstein, Proctor, Belcher, & Schurman, 1973, 1974; Boyer & Ro, 2007; Enns & DiLollo, 1997;

This research was supported by ESRC Grant RES-000-22-3087 to Angus Gellatly.

D. Guest (✉)
Faculty of Communication Scences,
University of Lugano (Università della Svizzera Italiana),
via Giuseppe Buffi 13,
6900 Lugano, Switzerland
e-mail: duncan.guest@usi.ch

A. Gellatly · M. Pilling
Oxford Brookes University,
Oxford, UK

Kahan & Mathis, 2002; Kahneman, 1968). Di Lollo et al. (2000) proposed that OSM is the result of an interaction between forward and backward projections of information through the visual system. Upon stimulus presentation, lower level cells are activated that have spatially local receptive fields and are tuned to relatively simple structures (e.g., orientation). As this activation is fed forward through the visual system, the stimulus gets represented at increasingly complex levels in cells with larger receptive fields. At higher levels of processing, competing pattern hypotheses are generated. To resolve this competition, information is projected back down the processing stream in order to compare activation at higher levels with activation at lower levels of processing. This resolution process requires several iterations, and if the stimulus remains constant over the number of iterations required to achieve a stable percept, it becomes consciously perceived. However, if there is a mismatch between activation at the different levels (activation at the lower level no longer supports the pattern hypothesis), the iterative cycle has to begin anew, the result being that only current sensory input enters conscious awareness. When the target and mask have a common offset, the *target-plus-mask* hypothesis can be matched to the persisting trace of the target plus mask, yielding a stable percept that is then consciously perceived. However, when the target offsets and the mask remains present, there is a mismatch between the target-plus-mask hypothesis and sensory input. Instead, there is now support for a *mask-only* hypothesis, and so the perception of the target plus mask is substituted by perception of the mask alone.

In support of the notion that OSM arises at the level of competing representations (e.g., pattern hypotheses) OSM does not appear greatly dependent on the physical or spatial relationship between target and mask. For example, it is not necessary that the contours of the target and mask are spatially close (Di Lollo et al., 2000; Enns & Di Lollo, 1997) or even, as long as attention is diffuse, that the target and mask are presented at the same location (Hirose & Osaka, 2009; Jiang & Chun, 2001a, b). Neither is it necessary that the target and mask resemble each other. For example, four dots that surround the target have been shown to mask a diamond missing a segment (Enns & Di Lollo, 1997), a circle missing a segment (Di Lollo et al., 2000), letters (Chen & Treisman, 2009; Enns, 2004; Neill, Hutchinson & Graves, 2002; Woodman & Luck, 2003), and even faces (Reiss & Hoffman, 2007).

Nevertheless, although physical resemblance between the target and mask does not appear to be necessary to produce robust OSM, Gellatly et al. (2006) found greater masking of target features when these target features matched those of the mask. Similarly, although OSM can be observed when target and mask do not share the same location, OSM arises only when the target and mask share the same depth plane or when the mask appears in front of the target (Kahan & Lichtman, 2006). Kahan and Lichtman also found that masking was greater if the mask moved toward the target location after target offset, as compared with when the mask and target initially shared the same location but, following target offset, the mask moved away from the target location. In combination, these experiments provide considerable support for the notion that spatial competition between target and mask can be important in OSM.

The purpose of the experiments presented here is to further examine spatial competition as a factor that influences OSM. Specifically, we examine whether the extent of overlap between target and mask influences OSM. Di Lollo et al. (2000) suggested that with a 4DM, the masking object should be considered as the surface produced by the modal completion of the square formed by the four dots. If the four dots were black on a white background, the 4DM would appear to be a white square (Kanisza, 1979). If the "target" is similarly defined as the surface produced by contours of the target, it can be seen how these two surfaces might compete at a given location. The extent of target–mask surface overlap may influence the extent of this competition and, hence, OSM.

Support for the notion that the notional surface of a 4DM is a relevant factor in OSM is found in the phenomenon of *object trimming* (Kahan & Enns, 2010; Kahan & Mathis, 2002). This is a phenomenon produced in OSM in which two dots arranged vertically and appearing at the side of a stimulus alter the percept resulting from that stimulus when these dots linger after the stimulus' offset. For example, when the dots of a two-dot trailing mask are presented above and below and to the left of a black diamond, the result is a percept in which the black diamond appears to have its left corner missing (Kahan & Mathis, 2002). The illusory surface, or contour, connecting the two dots, seems selectively to interfere with the processing of the portion of the stimulus that it flanks, resulting in an altered percept (Kahan & Enns, 2010).

In the experiments presented here, we examined whether the extent of overlap between the surfaces of a 4DM and the target would influence the degree of OSM. If spatial competition between target and mask does play a role in OSM, it might be expected that OSM would be greater the more the notional mask surface within the four dots overlapped with the target surface. If this were the case, an important question would be whether this notional mask surface selectively disrupts processing of the particular feature(s) of the target that it overlaps. The existence of object trimming certainly suggests that a mask surface can selectively interfere with specific features of the target object. However, if, as proposed by Di Lollo et al. (2000), OSM arises due to the representation of the *mask only* substituting for that of the *target-plus-mask* representation, it should not matter which target features the mask surface overlaps, since the representation of the target should be substituted in its entirety.

The sequence of experiments was as follows. In Experiment 1, we examined whether manipulating the overlap between target and mask surfaces would influence the extent of OSM and whether a 4DM would hinder processing of the specific target features it overlapped. Experiment 2 ruled out an alternative explanation of the results of Experiment 1. Finally, Experiment 3 examined whether OSM would be influenced by manipulating mask size and, hence, target–mask surface overlap, using masks that were smaller than and presented inside the target and ones that were larger than and surrounded the target.

Experiment 1a

The research program of which the following experiments are a part had approval from the Oxford Brookes University Research Ethics Committee.

Method

Participants Twenty-five participants from Oxford Brookes University (including D.G. and M.P.) took part in the experiment. Participants in this experiment (and in all the following experiments) were recruited from the Oxford Brookes participant panel and received either course credit or a small financial recompense (apart from D.G. and M.P., who participated in all the experiments except 1b).

Apparatus and stimuli A Pentium 4 was used for stimulus presentation and response registration. Participants sat 120 cm from the center of a 17-in. CRT monitor with a resolution of 1,024 × 768 and a refresh rate of 85 Hz. The experiment was written in MATLAB using the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997). Responses were made using the arrow keys on a standard keyboard.

The stimulus display contained four items presented either 2.65° above/below or to the left/right of the central fixation cross (0.24° × 0.24°). Stimuli were black outline squares with a width of 0.67° of visual angle. The black outline was 1 pixel in thickness. Each stimulus had a gap at the center of one side equal to 0.33°. Masks were 4DMs. The dots comprising the 4DM were black 0.05° × 0.05° squares. The notional square formed by the outer edges of the 4DM was 1.15° × 1.15° in size. The display background was white.

Design and procedure A two-factor within-subjects design was used, with mask duration and mask position as factors. The mask onset simultaneously with the display and either offset simultaneously with the display after 47 ms or was presented for 400 ms. The mask could be either centered over the target so that it surrounded it or offset so that the mask overlapped only half of the target. The mask could completely overlap the *critical* feature (i.e., the side with the gap in it), the side *adjacent* to the critical feature, or the side *opposite* the critical feature. Figure 1 shows a schematic of the trial sequence for the condition in which the mask overlapped the adjacent side. As can be seen, this resulted in partial overlap of the critical feature. On each trial, the mask was presented at one of the four stimulus locations and at one of the five mask positions (surround, critical, adjacent [× 2], opposite).

Participants completed two practice blocks (each block contained ten trials, one trial for each level of the design) and 35 experimental blocks. Trial order was randomized within each block. Each trial started with a small black fixation cross at the center of the screen, which stayed visible throughout the trial. After 500 ms, the array of stimuli was presented alongside the mask. On each trial, the mask was randomly presented at one of the four stimulus locations. The stimulus display offset after 47 ms, and the mask offset either simultaneously or after a further 353 ms. Participants had to indicate which side of the stimulus at the mask location had a gap in it. The gap could be present on any one of the four sides of the stimulus, and so the task was always a four-alternative forced choice. Responses were made using the arrow keys on the keyboard, and after responding, corrective feedback was provided. Every 70 trials, instructions were displayed during a break, and the average accuracy for the last 70 trials was displayed. Each experimental session took approximately 20 min.

Results and discussion

Figure 2a shows the percentage of correct responding at each mask position and mask duration, and Fig. 2b shows the size of the masking effect at each mask position (performance when the mask trailed the target subtracted from performance when the mask and target had a common offset). Data for the two adjacent mask positions were averaged in all analyses. A 2 (mask duration) × 4 (mask position) repeated measures ANOVA on the percentage of correct responses yielded a main effect of mask duration, $F(1, 24) = 78.00$, $MSE = .02$, $\eta_P^2 = .77$, $p < .001$, a main effect of mask position, $F(3, 72) = 14.87$, $MSE = .009$, $\eta_P^2 = .38$, $p < .001$, and a significant interaction between mask duration and mask position, $F(3, 72) = 3.65$, $MSE = .006$, $\eta_P^2 = .13$, $p = .016$. Comparisons showed that significantly greater masking was observed when the mask surrounded the target than when the mask overlapped the critical feature, $t(24) = 2.22$, $p = .036$, and the opposite feature, $t(24) = 3.12$, $p = .005$, but not when

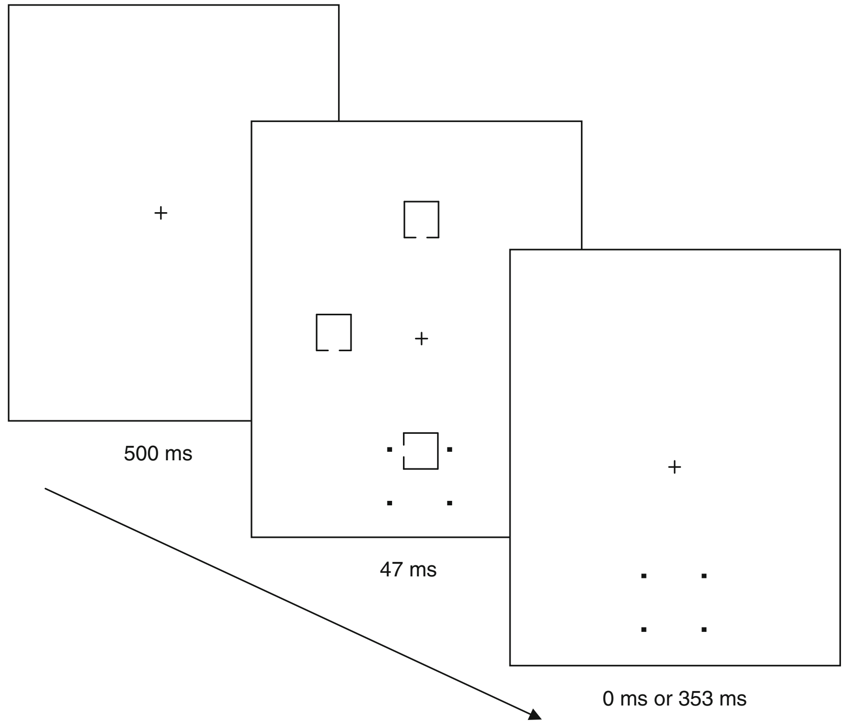

Fig. 1 Trial sequence in Experiment 1a. The mask was presented so that it surrounded the target or overlapped half of the target. In this example, the mask completely overlapped the side adjacent to the critical feature (the side with the gap in it)

the mask overlapped the adjacent feature, $t(24) = 1.02$, $p = .31$. Masking was more apparent when the mask overlapped the adjacent feature than when it overlapped the critical feature or the opposite feature, although these differences narrowly failed to reach significance, $t(24) = 1.96$, $p = .06$, and $t(24) = 1.95$, $p = .06$, respectively.

The central finding from Experiment 1a is that the amount of target–mask overlap has a large influence on masking. Masking tended to be greater when the mask surrounded the target and the overlap between the mask surface and target surface was greatest. Nevertheless, a considerable amount of masking was also observed when the mask overlapped the side of the target adjacent to the critical feature. This appears inconsistent with the notion that the extent of overlap between target and mask influences substitution masking, since the overlap between the target and mask was equivalent in the *critical*, *opposite*, and *adjacent* conditions. Since this somewhat anomalous aspect of the data needs to be accounted for, we first describe a subsidiary experiment performed to explain this anomaly before offering a proper interpretation of the findings from Experiment 1a.

A potential explanation for the finding of considerable masking when the mask surrounded either of the sides adjacent to the critical feature is that, in this condition, one of the four dots from the 4DM was presented close to the gap in the target (see Fig. 1). This may have resulted in some perceptual interaction between the stimulus and the dot and may have made it more difficult to correctly identify the gap in the target when the 4DM trailed the target offset. Since the dots of the 4DM were relatively large, participants would presumably have seen something black in the area where the gap was and could have misinterpreted this as indicating that there was full contour at this location. Such a misinterpretation may have been reinforced when the mask trailed the target, resulting in greater masking. Importantly this perceptual interaction should also have occurred when the target and mask had a common offset, potentially explaining why common-offset performance was significantly worse in the adjacent condition, as compared with the surround, critical feature, or opposite feature conditions, $t(24) = 6.14$, $p < .001$, [$t(24) = 3.51$, $p = .002$, and $t(24) = 3.78$, $p = .001$, respectively.

Experiment 1b

The purpose of Experiment 1b was to examine whether the large masking effect observed in the adjacent condition of Experiment 1a was, as hypothesized, due to the proximity of the dot to the gap in the square target. In Experiment 1b, a two-dot mask was used that was identical to the 4DM used in Experiment 1a, but with two dots instead of four. The two-dot mask was presented so that it bisected the target. Depending on which side of the target contained the critical feature (the gap), the mask could either bisect the critical feature or not. Only when the mask bisected the critical feature was one of the dots of the mask close to the gap. If the proximity of the dot to the gap interferes with the perception of the gap, a large masking effect should be observed when the mask bisects the critical feature, alongside a reduction in performance when

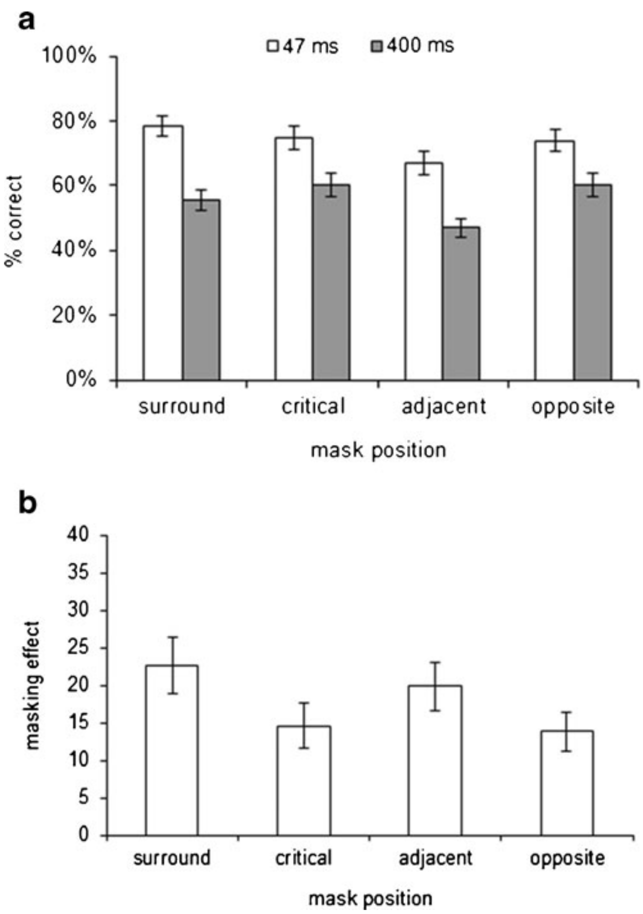

Fig. 2 Mean percentage of correct responses (**a**) and mean amount of masking (**b**) for masks presented at each mask position in Experiment 1a. Masks either surrounded the stimulus or completely overlapped the half of the target containing the *critical* feature (the side with the gap in), the side *adjacent* to the critical feature, or the side *opposite* the critical feature. Masks were displayed for either 47 ms (common offset with the stimulus array) or 400 ms (trailing mask). Standard errors are shown in both panels

the target and mask offset simultaneously (referred to hereafter as the *control condition*).

Method

Participants Twenty participants from Oxford Brookes University took part in the experiment.

Apparatus and stimuli The apparatus and stimuli were identical to those in Experiment 1a, except that a two-dot mask was used rather than a 4DM.

Design and procedure A two-factor within-subjects design was used, with mask duration and mask position as factors. The mask onset simultaneously with the display and either offset simultaneously with the display after 47 ms or was presented for 400 ms. The mask was always presented in a vertical orientation and centered over the target, bisecting it. The target was presented such that the mask either bisected the side with the critical feature and the side opposite the critical feature (*critical feature condition*) or bisected both sides adjacent to the critical feature (*not critical feature condition*).

Participants completed 2 practice blocks (each block contained four trials, one trial for each level of the design) and 40 experimental blocks. The order in which trials were presented was randomized within each block. The procedure was the same as that in Experiment 1a. Each experimental session took approximately 10 min.

Results and discussion

Figure 3 shows the percentage of correct responding at each mask position and mask duration, as well as the amount of masking observed at each mask position. A 2 (mask

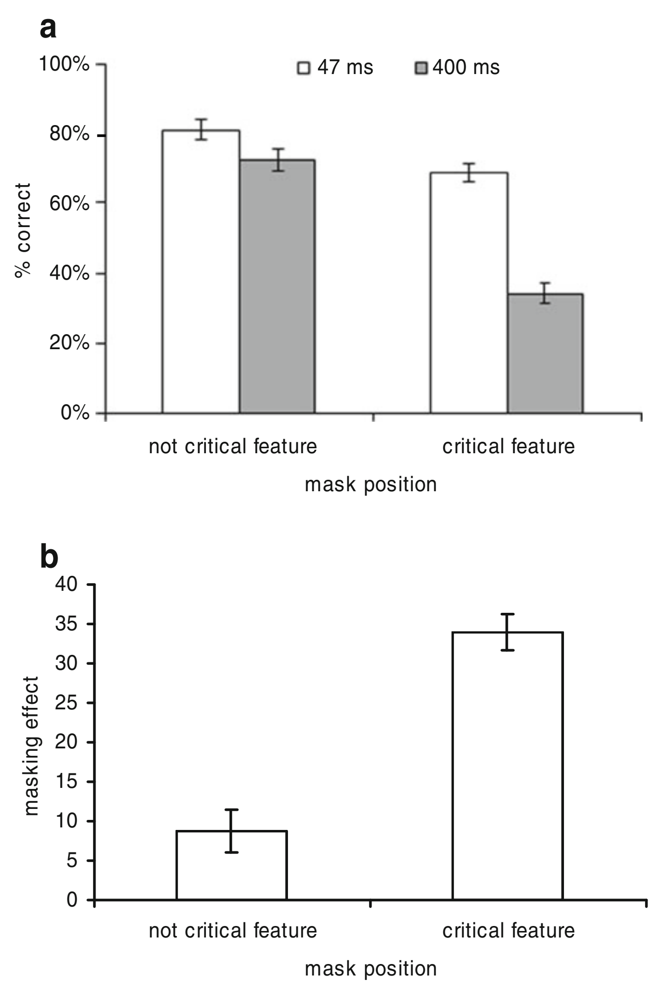

Fig. 3 Mean percentage of correct responses (**a**) and mean amount of masking (**b**) for masks presented at each mask position in Experiment 1b. Masks were presented so as to bisect the stimulus, with the critical feature either being bisected or not. Masks were displayed for either 47 ms (common offset with the stimulus array) or 400 ms (trailing mask). Standard errors are shown in both panels

duration) × 2 (mask position) repeated measures ANOVA on the percentage of correct responses yielded a main effect of mask duration, $F(1, 19) = 119.69$, $MSE = .008$, $\eta_P^2 = .86$, $p < .001$, a main effect of mask position, $F(1, 19) = 49.00$, $MSE = .026$, $\eta_P^2 = .72$, $p < .001$, and a significant interaction between mask duration and mask position, $F(1, 19) = 77.23$, $MSE = .004$, $\eta_P^2 = .80$, $p < .001$. In both the common-offset condition and the trailing mask condition, performance was significantly worse when the mask bisected the critical feature, $t(19) = 3.77$, $p = .001$, and $t(19) = 8.71$, $p < .001$, respectively. Comparisons showed that significantly greater masking was observed when the mask bisected the critical feature, as compared with when it did not, $t(19) = 8.79$, $p < .001$.

The results support the hypothesis that the proximity of one of the masking dots to the gap in the target interfered with the perception of the gap. Much greater masking was observed when the mask bisected the critical feature and one of the dots of the two-dot mask was close to the gap in the target. Moreover, performance in the control condition was also worse when the mask bisected the critical feature. This provides strong evidence that the large amount of masking observed in the adjacent condition in Experiment 1a was caused by lower level interference between the target and mask, rather than by object substitution.

An additional issue of interest is whether the two-dot mask disrupted processing of the target features it bisected. In object trimming, a two-dot mask presented adjacent to a target interferes with the perception of target features that are close to it, essentially eradicating them from perception (Kahan & Enns, 2010). If such disruption of feature processing occurred here when the mask bisected a target feature, then when the mask trailed the target, the resultant impression in the critical feature condition would be of a target with a gap in two of its sides, whereas the resultant impression when the mask did not bisect the critical feature would have been of a target with a gap in three of its sides. A consequence of this could be less accurate responding when the mask did not bisect the critical feature. However, if participants deduced that the mask interfered with the processing of the features it bisected, the opposite finding could result. That is, it would be easier to select the correct response in the condition where the mask did not bisect the critical feature because two out of three gaps could be accounted for by the mask position. These contrasting predictions make it difficult to interpret whether, as in object trimming, the two-dot mask interfered with the processing of target features. Moreover, even if such interference contributed to the masking effects observed, it could not explain why performance in the control condition differed depending on whether the mask bisected the critical feature or not.

Having accounted for the anomalous result from Experiment 1a whereby a mask surrounding the critical feature of the target produced less masking than did one surrounding a side adjacent to the critical feature, we can now return to considering the more substantive result from Experiment 1a—namely, that a mask surrounding the whole target produces more masking than does one surrounding the side containing the critical feature or the side opposite that. Crucially, since none of the four dots of the mask were close to the gap in the target in the *surround*, *critical*, and *opposite* conditions, these conditions were probably free of any effects of low-level perceptual interactions between target and mask. In both the critical and opposite conditions, masking was much reduced as compared with when the mask fully overlapped the stimulus, indicating that target–mask overlap does influence substitution masking.

Interestingly, the amount of masking was not dependent on whether the mask overlapped the critical target feature. Similar amounts of masking were observed when the mask overlapped the critical feature or the side opposite to it, despite the critical feature being left "uncovered" by the mask surface when the mask overlapped the opposite side. This suggests that OSM is not simply caused by the mask surface's occluding a specifically located feature of the target and preventing processing of it. Rather, it seems that the 4DM interferes with the representation of the whole target in the manner described in the original object substitution account of the phenomenon (DiLollo et al., 2000). That is, it seems that there is competition for consciousness between a mask object and a target object, not simply between mask features and target features at a given spatial location. We return to this in more depth in the General discussion section.

An alternative interpretation of Experiment 1a is that, rather than target–mask overlap determining the extent of masking, more masking is observed when the mask and the target have a common center (i.e., are concentric). According to Lleras and Moore (2003; see also Moore & Lleras, 2005), substitution masking occurs when a single object token is assigned to both the target and the mask. Object tokens are assigned to objects in order to track them over time across repeated samples of the environment. The feature information tied to an object token is continually updated as sensory information changes. Thus, according to the updating account, when the mask trails the target, the visual system interprets the mask as a continuation of the target plus mask, and thus the features of the target plus mask are overwritten by those of the mask alone. It is possible that, if the target and mask have a common center, the visual system is more likely to interpret them as belonging to the same object and, thus, assign them the same, single-object token. If so, more masking would be expected when the mask surrounded the whole target rather than half of the target, regardless of any overlap between the mask surface and the target surface. This prediction follows from consideration of how shapes are represented by the visual system relative to an axis of symmetry (see Attneave, 1971). Where the shape is regular, this axis passes through the center of the figure. The same figure will be perceived either as a square or as a diamond, depending on whether it is visually encoded relative to an axis of symmetry through the midpoints of two opposite sides or relative to a diagonal axis of symmetry. Because the figure can be perceived in either of two ways, it is said to display *bistability*. The visual system finds it economical to use the same axis or axes with a shared orientation to interpret as many figures in a scene as possible, with the result that, for example, a "flock" of ambiguous triangles will all appear to point in the same direction and to change the direction in which they point simultaneously (Attneave, 1971). This tendency of the visual system to encode shapes relative to the same axis or similar axes is strengthened by factors such as shared orientation, collinearity, and concentricity. Applied to Experiment 1a, this suggests that when the mask and target were concentric (surrounding mask), they were more likely to have been coded as a single object, relative to one of their shared axes of symmetry, than when they were not concentric. Experiment 2 therefore examined whether target and mask concentricity influences substitution masking.

Experiment 2

In Experiment 2, the location of the target within a 4DM was manipulated so that the target and mask were either concentric or not. In order to do this, a larger 4DM was used, and the size of the stimuli was reduced. This meant that it was immediately apparent when the target and mask did not share a common center.

Method

Participants Thirty-two participants from Oxford Brookes University (including D.G. and M.P.) took part in the experiment.

Apparatus and stimuli The apparatus and stimuli were identical to those in Experiment 1, except that in the stimulus display, the four stimuli were presented either 2.67° above/below or to the left/right of the central fixation cross. Since the location of the target within the 4DM was manipulated, a larger 1.69° 4DM was used, and the size of the stimuli was reduced to 0.45°.

Design and procedure A two-factor within-subjects design was used, with mask duration and mask position as factors. The mask onset simultaneously with the display and either offset simultaneously with the display after 35 ms or was presented for 400 ms. On any given trial, the mask was presented such that it had a common center with the target or was randomly shifted up/down/left/right by 0.53° so that its center differed from that of the target (see Fig. 4). Each block of trials comprised 16 trials, 8 trials with a mask with the same center as the target (4 with a common-offset mask

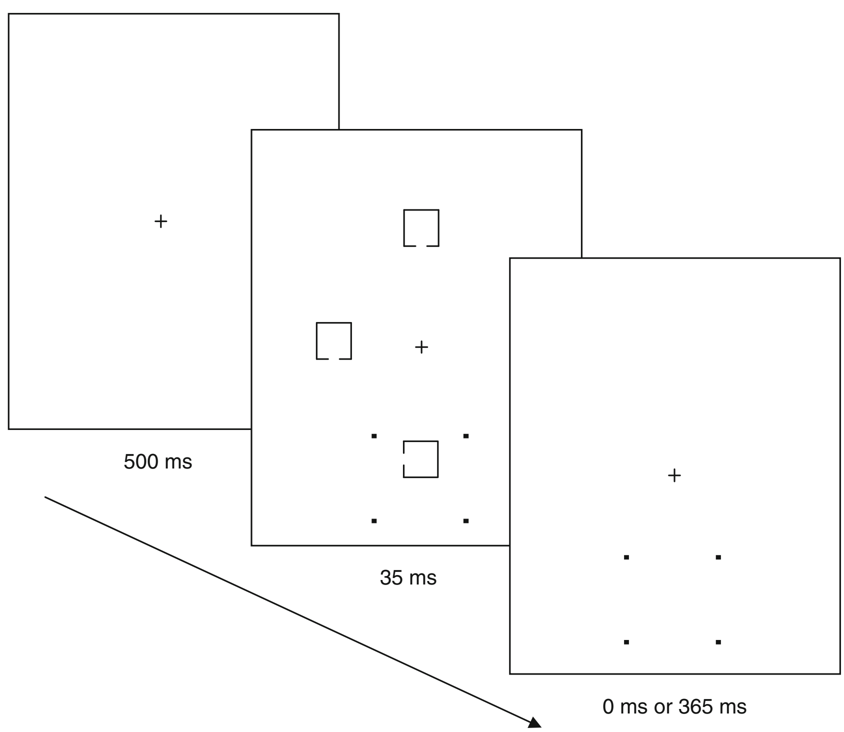

Fig. 4 Trial sequence in Experiment 2. The mask was presented so that it either shared the same center as the target or was shifted up/down/left/right by 0.53° so that the mask and target had different centers

and 4 with a trailing mask) and 8 trials in which the mask had a different center (2 trials for each up/down/left/right offset, 1 with a common-offset mask and 1 with a trailing mask). Participants completed two practice blocks and eight experimental blocks. The order in which the trials were presented was randomized within each block. The procedure was the same as that in Experiment 1. Each experimental session took approximately 20 min.

Results and discussion

Data for the four types of *different-center* trials were averaged in all analyses. Figure 5 shows the percentage of correct responding at each mask position and the amount of masking at each mask position. A 2 (mask duration) × 2 (common or different center) repeated measures ANOVA on the percentage of correct responses yielded a main effect of mask duration only, $F(1, 31) = 66.44$, $MSE = .017$, $\eta_P^2 = .68$, $p < .001$ (all other Fs < 1.01). Although the interaction between mask duration and manipulation of common center was not significant, there was a slight decrease in the amount of masking observed when the target and mask were not concentric. Assuming that the 2.54% difference in masking effect between the concentric and not concentric conditions was a small but reliable difference, the power of the experiment was only .16, suggesting that the experiment may not have had sufficient power to truly rule out an effect on OSM of whether or not the target and mask were concentric. However, since the effect of target–mask overlap in Experiment 1a was substantially larger than this 2.54% (a difference of 8% between masking in the surround condition, relative to the critical condition), it is doubtful whether the overlap effect was driven by the concentricity of target and mask.

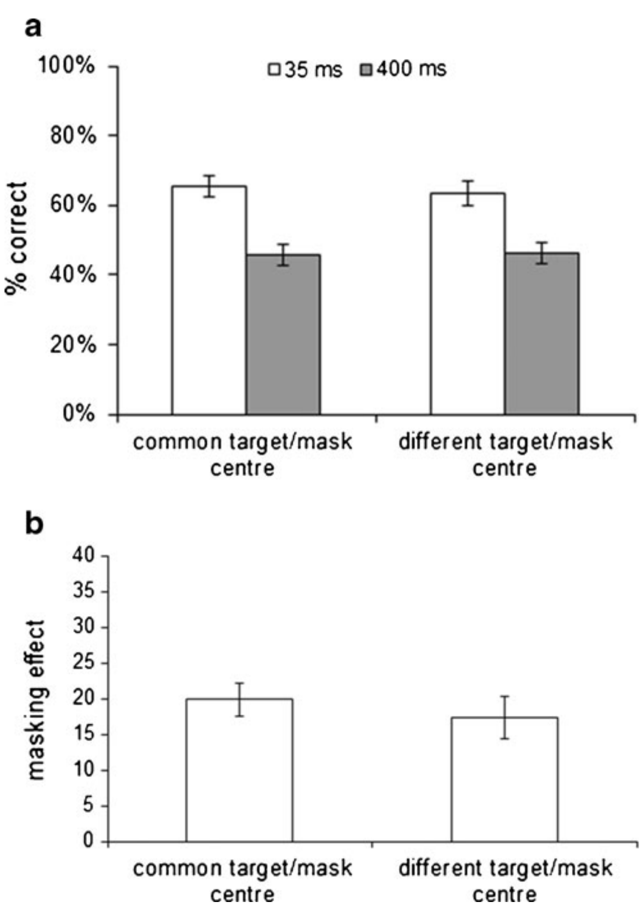

Fig. 5 Mean percentage of correct responses (**a**) and mean amount of masking (**b**) when target and mask had a common center and when they had a different center in Experiment 2. Masks were displayed for either 35 ms (common offset with the stimulus array) or 400 ms (trailing mask). Standard errors are shown in both panels

Since there were four types of spatial arrangement of mask and target in the different-center condition but only one type of spatial arrangement of mask and target in the common-center condition, the most common spatial arrangement of mask and target was for the target to appear at the center of the mask. This may have led to participants' developing a bias to attend to the center of the mask. This strategy could have artificially decreased accuracy in the different-center condition; however, accuracy levels were the same in both the common-center and different-center conditions, indicating that this does not seem to have happened.[1] Indeed, given that mask location and the positioning of the mask relative to the target stimulus changed from trial to trial, such a bias would take some time to be established and become effective. Moreover, on every trial, the stimuli were presented in exactly the same four locations (it was the mask location relative to the target stimulus that changed from trial to trial), so there was little reason for participants to attend anywhere other than these locations. As such, any effect of the spatial arrangement of the mask and target for the common-center condition being more frequent would have been relatively limited.

Although concentricity does not seem to be a relevant factor, it remains to be seen whether other means of influencing the relative encodings of target and mask may have an effect on masking. For example, might a square target be more strongly masked by a concentric 4DM perceived as a square than by one perceived as a diamond? For present purposes, however, the principal result of Experiment 1a can be attributed to a difference in target–mask overlap, rather than being a function of concentricity or its absence.

Experiment 3

Experiment 1 provides evidence that the amount of overlap between the surfaces of the target and mask is important in

[1] Our thanks to an anonymous reviewer for making this suggestion.

determining masking. However, OSM can be found using relatively small masks, such as a single dot (Lleras & Moore, 2003; Pilling & Gellatly, 2010). Moreover where mask size has been manipulated, it seems to have surprisingly little effect on the extent of OSM (Di Lollo et al., 2000; Enns & Di Lollo, 1997). The most extensive of these studies of mask size was performed by Di Lollo et al. (2000). Here, a range of mask sizes was used, with spatial separations between a 4DM and the contour of the target (a ring with a segment missing) taking values varying between 0.16° and 0.66°. Despite the extent of this range, there was little effect on the extent of masking and no evidence of any monotonic trend. However, the masks used in Di Lollo et al.'s (2000) experiment (and also those used by Enns & Di Lollo, 1997) were larger than and surrounded, the target. The fact that the mask surface completely overlapped the target surface in these experiments potentially explains the absence of any mask size effect. However, in one case, Di Lollo et al. (2000) also used a small 4DM where the four dots of the mask were clustered in the center of the stimulus. Counter to the notion that target–mask surface overlap influences masking, this small mask generally produced similar levels of OSM as masks that surrounded the target, except for when the set size was largest, where the small mask appeared to be less effective. Nevertheless, a potential problem with the Di Lollo et al. (2000) study is that the duration of the stimulus display was exceptionally brief (10 ms). Simultaneous onset of two stimuli, one of which is very brief, results in smooth apparent motion toward the trailing stimulus Di Lollo et al. (1993). This apparent motion is even more pronounced when the trailing stimulus has a long duration Di Lollo et al. (1993). The short target duration followed by the long mask duration (160 and 320 ms) in the Di Lollo et al. (2000) study may, therefore, have created the impression of the target (a circle with a gap in it) transforming into the mask (a 4DM). Lleras and Moore (2003) have previously demonstrated that such apparent transformation produces OSM (see also Pilling & Gellatly, 2010). Moreover, this is the case even though the target and mask are dissimilar (a circle with a gap (Lleras & Moore, 2003), a square with a gap, or a triangle (Pilling & Gellatly, 2010) transforming into a small dot). If the levels of masking observed by Di Lollo et al. (2000) were caused in part by this apparent transformation, this may have rendered the experiment insensitive to any manipulations of mask size. In Experiment 3, we reinvestigated the effects of mask size manipulations, using longer target durations (47 ms). Four different mask sizes were used; two masks were larger than and surrounded the target, and two masks were smaller than and presented inside the target.

Method

Participants Twenty-four participants from Oxford Brookes University (including D.G. and M.P.) took part in the experiment.

Apparatus and stimuli The apparatus and stimuli were the same as those in Experiment 1, with the following exceptions. Displays consisted of eight stimuli presented in an evenly distributed circular array (radius to the midpoints of stimuli was 1.74°) surrounding a central fixation cross. Masks were 4DMs of four sizes: 0.13°, 0.40°, 0.94°, or 1.20° (see Fig. 6a). Two masks were smaller than the target stimulus, and two were larger than the target stimulus. The distance between mask contour and target contour was the same for the smallest and largest masks (0.13°) and for the second smallest and second largest masks (0.27°). The display background was white. An example of the display is shown in Fig. 6b.

Design and procedure A two-factor within-subjects design was used, with mask duration and mask size as the factors. There were two levels of mask duration (47 or 400 ms) and four levels of mask size. Participants completed 2 eight-trial practice blocks (each block contained 1 trial for each level of the design) and 30 experimental blocks, with a break every 60 trials. Trials were randomized within each block. Each experimental session took approximately 15 min.

Results and discussion

The percentage of correct responses and standard error for each mask size and mask duration are shown in Fig. 7a. A 2 (mask duration) × 4 (mask size) repeated measures ANOVA was performed on the percentage correct data. As was expected, performance was worse when the mask trailed the stimulus, resulting in a significant main effect of mask duration, $F(1, 23) = 227.75$, $MSE = .011$, $\eta_P^2 = .91$, $p < .001$. Performance also declined with increasing mask size, resulting in a significant main effect of mask size, $F(3, 69) = 27.26$, $MSE = .018$, $\eta_P^2 = .54$, $p < .001$. The interaction between mask size and mask duration was also significant, $F(3, 69) = 3.29$, $MSE = .008$, $\eta_P^2 = .16$, $p = .026$, with much less masking observed for the smallest mask (see Fig. 7b). Further analysis revealed that the amount of masking produced by the smallest mask was much less than the masking produced by the 0.40° size mask, $t(23) = 2.94$, $p = .005$, the 0.94° size mask, $t(23) = 2.87$, $p = .009$, and the 1.20° size mask, $t(23) = 2.17$, $p = .041$. There were no other significant differences in the amount of masking produced by the different mask sizes.

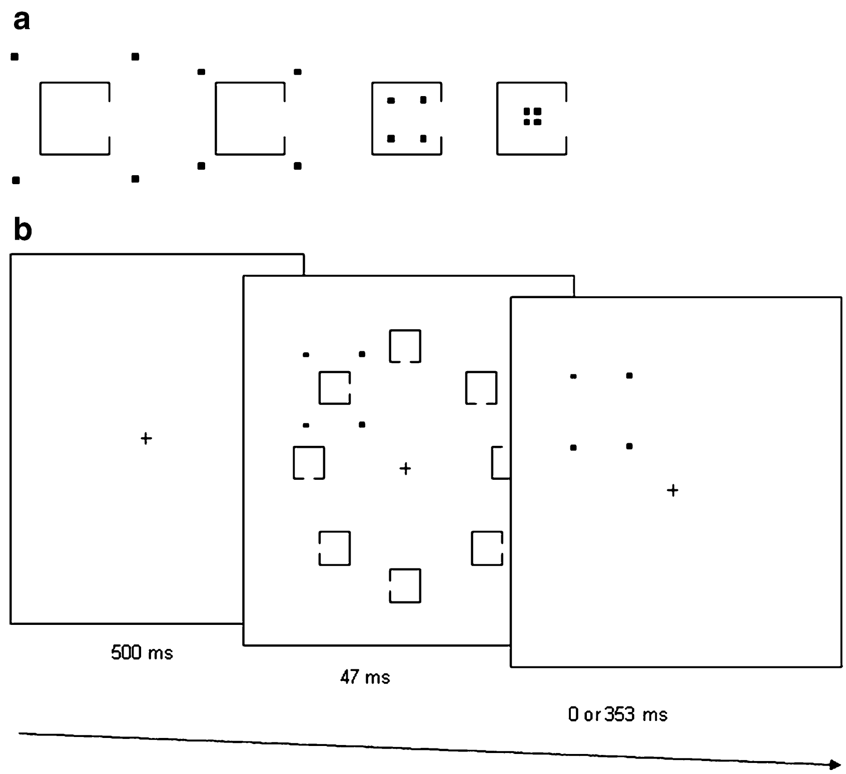

Fig. 6 **a** Different-sized masks used in Experiment 3. **b** Trial structure in Experiment 3. A 500-ms fixation cross was presented, followed by the onset of the stimulus display and the mask. The stimulus display offset after 47 ms, and the mask offset after either 47 or 400 ms

Consistent with previous studies, manipulating mask size had no effect on the amount of substitution masking for masks that were larger than the target (Di Lollo et al., 2000; Enns & Di Lollo, 1997). However, masking was reduced for a small mask presented inside the target, contrasting with Di Lollo et al. (2000), who found similar levels of masking for a small inside mask and larger outside masks. This may be because the short target duration in the Di Lollo et al. (2000) study created the impression of the target transforming into the mask, which is known to produce OSM (Lleras & Moore, 2003). The longer target duration used here would have prevented such an apparent transformation, providing a better assessment of the effects of reducing mask size.

Although somewhat consistent with the notion that a reduction in mask size should reduce levels of OSM due to reduced mask–target overlap, a problem with this interpretation is that relative to the outside masks, OSM was reduced only for the smallest inside mask. It is possible that the amount of target–mask overlap for the largest inside mask (0.40° mask) was still enough to produce maximal spatial competition between target and mask. Alternatively, it may be that mask size per se is critical in determining the extent of OSM irrespective of target–mask overlap. Very small masks may not compete for and hold attention to the extent of larger masks—perhaps because they segregate better from the target (see Fig. 6a)—allowing attention to be more easily directed to the fading trace of the target and, thus, reducing masking.

Mask size effects may also have been influenced by lower level visual processes—for example, metacontrast masking, which occurs when the mask surrounds but does not overlap with the target contour. Since classical metacontrast masking is sensitive to distance between target and mask contour (Breitmeyer, 1984), previous studies showing no effect of manipulating the size of 4DMs have argued that this differentiates OSM from metacontrast masking (Di Lollo et al., 2000; Enns & Di Lollo, 1997). Nevertheless, contour proximity is not a necessary condition for metacontrast masking, particularly when the task involves discrimination of target contour (Breitmeyer et al., 1976; Breitmeyer et al., 1974; Breitmeyer & Ogmen, 2000), raising the possibility that the small 4DM produced the least masking because it had the least contour, whereas the contour of the other-sized masks was similar in size to that of the target. There are several reasons to believe that this was not the case. First, the use of a 4DM is supposed to remove the possibility of metacontrast masking, because the visible contour of the four dots is very small and the larger square formed by the 4DM has a negligible amount of actual contour with which to interfere with the lower level visual processes (Di Lollo et al., 2000). Indeed, the invisible contour of the square formed by the global completion of the 4DM is itself a result of higher level visual processes, whereas metacontrast masking is assumed to be caused by lower level visual processes. In addition, metacontrast masking, unlike OSM, is typically not observed when target and mask onset simultaneously (e.g.,

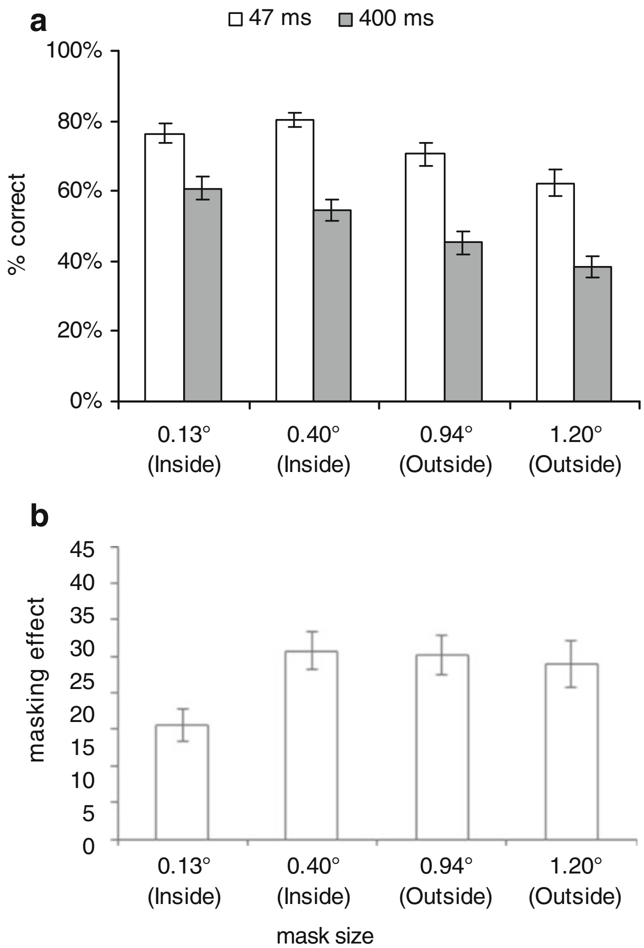

Fig. 7 Mean percentage of correct responses (**a**) and mean amount of masking (**b**) for masks of different size in Experiment 3. Masks were small and presented inside the target or were large and presented outside the target. Masks were displayed for either 47 ms (common offset with the stimulus array) or 400 ms (trailing mask). Standard errors are shown in both panels

Bernstein et al., 1973, 1974; Boyer & Ro, 2007; Kahneman, 1968), whereas large amounts of masking were observed in Experiment 3 under these conditions. Because of this, we conclude that the smallest mask in Experiment 3 produced the least masking due to its reduced overlap with the target and, hence, because of the spatial competition between the surfaces of the target and mask items.

General discussion

In the last 10 years, much research has examined the conditions under which object substitution masking can be produced. A number of studies have demonstrated that OSM can be produced even when the mask is not presented at the target location, as long as attention is diffuse (Jiang & Chun, 2001a, b) or the mask and target appear to move along the same trajectory (Moore & Lleras, 2005). Indeed, it is even possible to produce OSM without a mask when other distracting objects remain present after the offset of the target (Luiga, Gellatly, & Bachman, 2010). Despite this, it does appear that one of the factors influencing OSM is the extent to which there is spatial competition between the target and mask. For example, Kahan and Lichtman (2006) found that when target and mask are presented in the same location and it appears that the mask is either in the same depth plane as or is in front of the target, robust OSM is produced. However, if it appears that the mask is behind the target, no OSM is observed.

The experiments reported here further examined the influence of spatial competition on OSM. The premise of the experiments was that the area inside the mask be considered as a surface produced by the modal completion of its contours, so that a 4DM composed of four dots is perceived as a filled square (Di Lollo et al., 2000). If the target object is defined similarly, spatial competition between a target and mask may be determined by the extent to which the surfaces of the target and mask overlap. Experiments 1–3 examined whether the extent of overlap between target and mask surfaces influences OSM. In Experiment 1, the 4DM either surrounded the target or overlapped half of the target, with greater masking produced when the mask surrounded the target. Experiment 2 indicated that the effect of target–mask surface overlap in Experiment 1 was not due to the target and mask being concentric in the surround mask condition, but not in the conditions where it overlapped half of the target. In Experiment 3, we tried to reconcile the effect of target–mask surface overlap observed in Experiment 1 with a previous finding that a small mask presented at the center of a target produced as much masking as did larger masks that surrounded the target, even though, in the latter case, target–mask surface overlap was considerably greater (Di Lollo et al., 2000). In the Di Lollo et al. (2000) experiment, the target was presented for a very brief duration, potentially creating apparent transformation between the mask and the target, which is known to produce OSM (Lleras & Moore, 2003). To minimize apparent transformation in Experiment 3, targets were presented for a longer duration. Under such circumstances, we found much reduced masking when the mask was small and presented at the center of the target, as compared with when it was larger and target–mask surface overlap was considerably greater. Combined, Experiments 1–3 provide support for the notion that spatial competition between target and mask plays a role in determining OSM.

An intuitive explanation for the finding that increasing surface overlap between the target and mask increases OSM is that the mask surface occludes part of the target surface, interfering with the processing of target information in the occluded area. However, Experiment 1 provides strong counterevidence to this hypothesis. In Experiment 1, OSM was reduced if the mask overlapped only half of the target, but it did not matter whether or not the half of the target that was overlapped contained the critical feature to be identified (a gap). This suggests that the effect of target–mask surface overlap on OSM is not due to occlusion of specific target features by the mask. Rather, it appears that the effect of target–mask surface overlap is the result of competition at the level of object representations, either between competing pattern hypotheses (Di Lollo et al., 2000) and interfering object files (Kahan & Lichtman, 2006) or because the target, or target and mask, is seen as morphing into the mask (Lleras & Moore, 2003). In all these interpretations, when the mask outlasts the target, competition tends to be resolved in favor of the mask, with the result that the representation of the target or target plus mask gets substituted by the representation of the mask only. The fact that in Experiment 1 the mask appeared to interfere with the processing of the whole object supports the notion that in these conditions, substitution occurs due to processes at the object level and not, as has been found under different stimulus arrangements (e.g., Gellatly et al., 2006), at the featural level.

This finding is surprising in that in contrasts with the phenomenon of object trimming whereby a two-dot mask presented on one side of a stimulus interferes with the processing of object features on that side of the object when it trails the target (Kahan & Enns, 2010; Kahan & Mathis, 2002). It is as if the contour connecting the two dots simply occludes nearby feature information, completely altering the appearance of the object. One key difference between the studies reported here and the object-trimming studies is that those reported here used 4DMs and not two-dot masks, as in most of the experiments of Kahan and colleagues (Kahan & Enns, 2010; Kahan & Mathis, 2002). Indeed, in Experiment 2 of Kahan and Enns, masking by two dots was reduced by the addition of a curved bar on the far side of the two dots from the target. By grouping with them, the curved bar perceptually "pulled" the two dots away from the target contour and reduced trimming. Perhaps then, by grouping together and generating the surface of a virtual square, the dots of our 4DMs entered into competition with the object-level representation of the target, rather than with a feature-level representation. Another key difference is that in the majority of cases, the contours of the masks in Experiments 1–3 were not presented in close proximity to one particular side of the target, potentially diminishing the impact of object trimming. In sum, it is clear that the effect of target–mask surface overlap is not due to an object-trimming-like process in which processing of specific features is impaired by the overlapping mask.

What other processes may underlie the effect of target–mask overlap on OSM? According to Di Lollo et al. (2000), substitution masking is the result of a process in which the perceptual system chooses which, out of several pattern hypotheses generated from sensory input, best reflects the state of the world. This resolution process involves ongoing communication between lower and higher processing levels whereby pattern hypotheses are compared with incoming sensory information. A mismatch between a pattern hypothesis and incoming sensory information means that a new pattern hypothesis has to be generated. When the mask trails the target, a mismatch occurs, the comparison process begins anew, and the perception of the mask only substitutes the perception of the target plus mask. Although Di Lollo et al. (2000) said little about the impact of spatial competition between target and mask, it is possible that when the masks do not completely overlap the target, the fading trace of the target can be better attended. According to Di Lollo et al. (2000), when an object is attended, fewer feedback sweeps are required for it to enter conscious awareness, and so attending to an object reduces OSM. This would certainly explain why manipulating target–mask surface overlap influences masking. However, it is more difficult to reconcile with the finding that it does not matter whether or not the mask overlaps the critical feature, simply that it overlaps some portion of the target, since it might be expected that attending to the critical feature would reduce masking.

Another related interpretation of substitution masking is that it occurs when the visual system fails to assign different object tokens to the target and the mask (Enns, Lleras, & Moore, 2009; Hein & Moore, 2010; Lleras & Moore, 2003; Moore & Lleras, 2005). Object tokens are assigned to objects in the environment in order to maintain a stable representation of these objects as the information sampled from the external world changes. Since sensory information about an object constantly changes, feature information tied to each token is continually updated. If the target and mask are mistakenly assigned the same object token, the result is that information about the trailing mask will overwrite any information about the target, producing masking. In contrast, if the target and mask are assigned separate object tokens, the information about the trailing mask will overwrite information associated with the "mask" token only, and so no masking will occur.

At first blush, it is difficult to see how the object-updating hypothesis could explain the results of this series of experiments. However, consider the problem faced by the visual system in the experiments above. When the target and mask are presented simultaneously but the target offsets quickly and the mask trails the target, the perceptual system has to decide whether the target and the mask are in fact the same object. To do this, the perceptual system could use spatial information, such as whether the target and mask are presented in the same location. However, in cases where target and mask are presented in the same location, spatial information by itself has limited use. The perceptual system may therefore combine this spatial information with information conveyed by any changes in the scene. When the mask is larger than the target and surrounds it, then when the target offsets and the mask remains onscreen, there are no changes in the global contour of the shape defined by the target and mask. The perceptual system may therefore conclude that the pre- and postchange objects are in fact the same and assign them the same object token. However, when the mask overlaps half of the target (Experiment 1) or is smaller than the target (Experiment 3), target offset results in a change in the global contour of the shape defined by the target and the mask. If this change is great enough, the perceptual system may interpret it as evidence that the mask and target are different objects and assign them different object tokens, reducing masking. The object-updating account can, therefore, explain the results of Experiments 1 and 3 without the need to infer that occlusion of the target by the mask surface hinders its processing.

The object-updating account of OSM can, therefore, provide a coherent account of the findings of Experiments 1–3. The degree of target–mask overlap influences the extent of object substitution masking, not because the mask occludes the target features, but because the greater the overlap, the less the contour of the combined target and mask object changes when the mask trails the target. Nevertheless, although feasible in the context of the experiments presented here, this account cannot explain why Kahan and Lichtman (2006) found that OSM was eliminated when the mask appeared as though it was behind the target, even though both the target and mask were in the same spatial location. Since this effect of spatial competition cannot be explained via the object-updating account, Kahan and Lichtman suggested that there are two processes involved in OSM, spatial competition and object updating, and this has been supported by subsequent research (Guest, Gellatly, & Pilling, 2011; Kahan & Enns, 2010; Pilling & Gellatly, 2010). Further work will be required to determine whether the effect of mask–target surface overlap is primarily dependent on one or the other of these processes.

References

Attneave, F. (1971). Multistability in perception. In R. Held (Ed.), *Image, object and illusion* (pp. 91–99). San Francisco: Freeman.

Bernstein, I. H., Proctor, R. W., Belcher, J., & Schurman, D. L. (1973). Metacontrast and brightness discrimination. *Perception & Psychophysics, 14,* 293–297.

Bernstein, I. H., Proctor, R. W., Belcher, J., & Schurman, D. L. (1974). An analysis of U-shaped metacontrast. *Perception & Psychophysics, 16,* 329–336.

Boyer, J., & Ro, T. (2007). Attention modulates metacontrast masking. *Cognition, 104,* 135–149.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10,* 433–436.

Breitmeyer, B. G. (1984). Visual masking: An integrative approach. New York: Oxford University Press.

Breitmeyer, B., Battaglia, F., & Weber, C. (1976). U-shaped backward contour masking during stroboscopic motion. *Journal of Experimental Psychology: Human Perception and Performance, 2,* 167–173.

Breitmeyer, B. G., Love, R., & Wepman, B. (1974). Contour masking during stroboscopic motion and metacontrast. *Vision Research, 14,* 1451–1456.

Breitmeyer, B. G., & Ogmen, H. (2000). Recent models and findings in visual backward masking: A comparison, review, and update. *Perception & Psychophysics, 62,* 1572–1595.

Chen, Z., & Treisman, A. (2009). Implicit perception and level of processing in object-substitution masking. *Psychological Science, 20,* 560–567.

Di Lollo, V., Bischof, W. F., & Dixon, P. (1993). Stimulus-onset asynchrony is not necessary of motion perception or metacontrast masking. *Psychological Science, 3,* 260–263.

Di Lollo, V., Enns, J. T., & Rensink, R. A. (2000). Competition for consciousness among visual events: The psychophysics of reentrant visual processes. *Journal of Experimental Psychology: General, 129,* 481–507.

Dux, P. E., Visser, T. A. W., Goodhew, S. C., & Lipp, O. V. (2010). Delayed re-entrant processing impairs visual awareness: An object substitution masking study. *Psychological Science, 21,* 1242–1247.

Enns, J. T. (2004). Object substitution and its relation to other forms of visual masking. *Vision Research, 44,* 1321–1331.

Enns, J. T., & Di Lollo, V. (1997). Object substitution: A new form of masking in unattended visual locations. *Psychological Science, 8,* 135–139.

Enns, J. T., & Di Lollo, V. (2000). What's new in visual masking? *Trends in Cognitive Sciences, 4,* 345–352.

Enns, J. T., Lleras, A., & Moore, C. M. (2009). Object updating: A force for perceptual continuity and scene stability in human vision. In R. Nijhawan (Ed.), *Problems of space and time in perception and action* (pp. 503–520). Cambridge: Cambridge University Press.

Gellatly, A. R. H., Pilling, M., Carter, W., & Guest, D. (2010). How does target duration affect object substitution masking? *Journal of Experimental Psychology: Human Perception and Performance, 36,* 1267–1279.

Gellatly, A. R. H., Pilling, M., Cole, G. G., & Skarratt, P. (2006). What is being masked in object substitution masking? *Journal of Experimental Psychology: Human Perception and Performance, 32,* 1422–1435.

Guest, D., Gellatly, A. R. H., & Pilling, M. (2011). *Why is OSM reduced when target duration is increased, loading into VSTM or individuation of mask and target?* Manuscript submitted for publication.

Hein, E., & Moore, C. M. (2010). Unmasking the standing wave of invisibility: An account in terms of object-mediated representational updating. *Attention, Perception, & Psychophysics, 72,* 398–408.

Hirose, N., & Osaka, N. (2009). Object substitution masking induced by illusory masks: Evidence for higher object-level locus of interference. *Journal of Experimental Psychology: Human Perception and Performance, 35,* 931–938.

Jiang, Y. H., & Chun, M. M. (2001a). Asymmetric object substitution masking. *Journal of Experimental Psychology: Human Perception and Performance, 27,* 895–918.

Jiang, Y. H., & Chun, M. M. (2001b). The spatial gradient of visual masking by object substitution. *Vision Research, 41,* 3121–3131.

Kahan, T. A., & Enns, J. T. (2010). Object trimming: When masking dots alter rather than replace target representations. *Journal of Experimental Psychology: Human Perception and Performance, 36,* 88–102.

Kahan, T. A., & Lichtman, A. S. (2006). Looking at object-substitution masking in depth and motion: Toward a two-object theory of object substitution. *Perception & Psychophysics, 68,* 437–446.

Kahan, T. A., & Mathis, K. M. (2002). Gestalt grouping and common onset masking. *Perception & Psychophysics, 64,* 1248–1259.

Kahneman, D. (1968). Methods, findings, and theory in studies of visual masking. *Psychological Bulletin, 70,* 404–425.

Kanizsa, G. (1979). *Organization in vision: Essays on Gestalt perception*. New York: Praeger.

Lleras, A., & Moore, C. M. (2003). When the target becomes a mask: Using apparent motion to isolate the object component of object-substitution masking. *Journal of Experimental Psychology: Human Perception and Performance, 29,* 106–120.

Luiga, I., Gellatly, A. R. H., & Bachmann, T. (2010). Delayed offset of distracters masks a local target. *Acta Psychologica, 134,* 344–352.

Moore, C. M., & Lleras, A. (2005). On the role of object representations in object substitution masking. *Journal of Experimental Psychology: Human Perception and Performance, 31,* 1171–1180.

Neill, W. T., Hutchinson, K. A., & Graves, D. F. (2002). Masking by object substitution: Dissociation of masking and cuing effects. *Journal of Experimental Psychology: Human Perception and Performance, 28,* 682–694.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10,* 437–442.

Pilling, M., & Gellatly, A. R. H. (2010). Object substitution masking and the object updating hypothesis. *Psychonomic Bulletin & Review, 17,* 737–742.

Reiss, J. E., & Hoffman, J. E. (2007). Disruption of early face recognition processes by object substitution masking. *Visual Cognition, 15,* 789–798.

Turvey, M. T. (1973). On peripheral and central processes in vision: Inferences from an information-processing analysis of masking with patterned stimuli. *Psychological Review, 81,* 1–52.

Woodman, G. F., & Luck, S. J. (2003). Dissociations among attention, perception and awareness during object-substitution masking. *Psychological Science, 14,* 605–611.